



**You have downloaded a document from
RE-BUS
repository of the University of Silesia in Katowice**

Title: Sieciowe rozpowszechnianie baz danych dostarczanych na nośnikach optycznych

Author: Andrzej Koziara

Citation style: Koziara Andrzej. (1999). Sieciowe rozpowszechnianie baz danych dostarczanych na nośnikach optycznych. "Biuletyn EBIB" (Nr 8, (1999)).



Uznanie autorstwa - Licencja ta pozwala na kopiowanie, zmienianie, rozprowadzanie, przedstawianie i wykonywanie utworu jedynie pod warunkiem oznaczenia autorstwa.



UNIwersYTET ŚLĄSKI
W KATOWICACH



Biblioteka
Uniwersytetu Śląskiego



Ministerstwo Nauki
i Szkolnictwa Wyższego



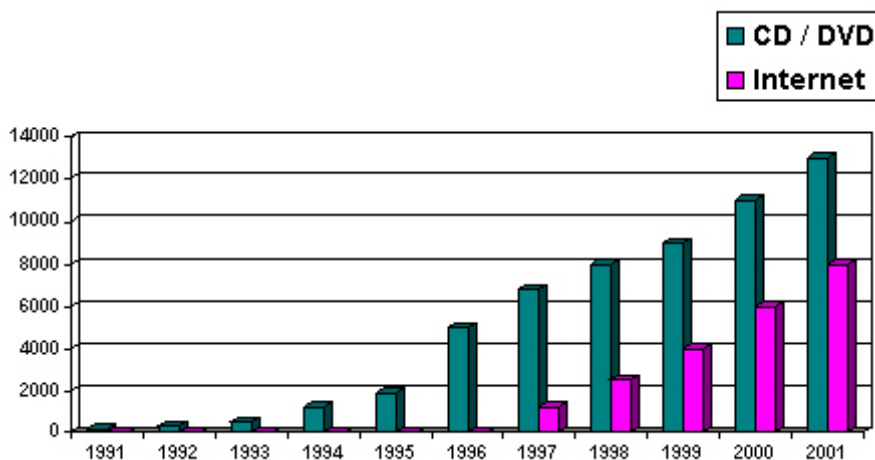
Andrzej Koziara
Biblioteka Główna Uniwersytetu Śląskiego

Sieciowe rozpowszechnianie baz danych dostarczanych na nośnikach optycznych

Lata dziewięćdziesiąte dwudziestego wieku bezsprzecznie możemy nazwać erą krzemu i plastiku. Rozwój technologii wytwarzania pierwszego z nich zadecydował o coraz większych możliwościach obliczeniowych współczesnych mikrokomputerów. Dotyczy to zarówno mocy obliczeniowej procesorów jak i coraz większych modułów pamięci operacyjnej. Towarzyszy temu również gwałtowny postęp technologiczny w branży produkcji pamięci masowych. W ciągu ostatnich lat praktycznie każdy rok przynosi dwu- a nawet trzykrotny wzrost pojemności dysków twardych bez zmiany ich ceny.

Zjawiska wymienione powyżej są dopełnieniem rozwoju innej branży technologicznej, powiązanej z produkcją i wytwarzaniem nośników optycznych. W ciągu ostatnich dziesięciu lat i w tym zakresie dokonano ogromnego skoku technologicznego. Postęp w tej dziedzinie następuje w dwóch nurtach: rozwoju nośników optycznych oraz obsługujących ich czytników. Pierwszy spowodował upowszechnienie się DVD-ROM (o teoretycznej pojemności nawet do 17 GB - przy zapisie dwustronnym), uzupełniających stosowane od lat płyty CD-ROM o standardowej pojemności 650 MB. Drugi zmierza w kierunku tworzenia urządzeń czytających płyty optyczne szybciej i z mniejszą ilością błędów.

Prace te napędzane są dodatkowo gwałtownym wzrostem ilości produkowanych tytułów, rozpowszechnianych za pomocą tej techniki. Mimo ogólnej opinii o możliwości zaniku tej technologii - po dokładnym przeanalizowaniu rynku - zauważamy, że publikacje te rozwijają się w takim samym tempie jak te, dla których jedynym medium jest światowy Internet.



Jak wynika z danych zbieranych przez firmę Watrlow New Media Information (dawniej TFLP), historia i prognozy wykorzystania tego nośnika są nadzwyczaj optymistyczne. Opublikowane dane pokazują, że rozwój ten jest w pełni zrównoważony i pozwala na spokojnie inwestowanie w wymienione wyżej technologie.

W związku z tą sytuacją niezwykle popularne stały się systemy sieciowego rozpowszechniania tak opublikowanych baz danych w sieciach lokalnych i rozległych. Szczególnie cenne są te drugie rozwiązania, gdyż zapewniają łatwy dostęp do danych na bazie lokalnych (miejskich, uczelnianych czy środowiskowych), niezbyt mocnych serwerów sieciowych, administrowanych przez zespoły pracowników informacji naukowej oraz zespoły informatyczne bibliotek naukowych.

Wśród tych rozwiązań możemy wyróżnić dwie grupy systemów:

- I. Uniwersalne - zapewniające obsługę praktycznie wszystkich baz danych dostarczanych przez ich producentów:
 - o Generacja pierwsza - sieciowo udostępniane czytniki CD-ROM
 - o Generacja druga - sieciowo udostępniane czytniki CD-ROM przy pomocy specjalistycznego oprogramowania adresującego napędy
 - o Generacja trzecia - bazy danych przeładowywane na dyski twarde, dostępne przy pomocy specjalistycznego oprogramowania klienta logującego użytkownika do serwera baz danych
 - o Generacja czwarta - bazy danych rozpowszechniane po przeładowaniu na dyski twarde z logowaniem do systemu poprzez przeglądarkę internetową (do przeszukiwania bazy danych wykorzystywane jest oprogramowanie fabrycznie dostarczone przez dowolnego jej producenta)
- II. Firmowe - przeznaczone do sieciowego rozpowszechniania baz danych w strukturze danych zgodnej z firmowym formatem :
 - o Silver Platter - serwery ERL z oprogramowaniem WinSpirs (równocześnie obsługujące tryb Workstation i client-server) i WebSpirs (oprogramowanie przekazujące wyniki przeszukiwań w formacie HTML)
 - o OVID - serwery danych obsługujące dostęp poprzez interface WWW, technologię client Telnet oraz w wersjach lokalnych
 - o Bell+Howell (dawniej UMI) technologia ProQuest serwis świadczony z serwera firmowego oraz programy do obsługi lokalnych baz danych (m.in. technologia sprzęgająca bazy indeksowe z bazami pełnotekstowymi dla stanowisk lokalnych)
 - o Dialog - technologia obsługująca usługi dostępne przez przeglądarki WWW oraz obsługująca bazy lokalnie

W związku z dużym rozmiarem materiału niezbędnego do przekazania czytelnikowi, omówienie tych zagadnień i systemów zostanie dokonane w dwóch częściach. W pierwszej części, zamieszczonej w tym numerze, przedstawiona zostanie historia systemów uniwersalnych oraz najbardziej popularny w Polsce system, dostarczany bezpłatnie z bazami danych przez firmę Silver Platter ERL. Należy przy tym zaznaczyć, że o popularności tego systemu decyduje oferta bibliograficznych baz danych firmy Silver Platter, obejmująca ponad 350 tytułów.

Systemy uniwersalne

Generacja pierwsza

Systemy opierały się o proste serwery plików, najczęściej pracujące w oparciu o oprogramowanie Novell Netware. W tak przygotowanym systemie przydzielano pewną ilość fizycznych napędów CD-ROM wraz z logicznymi odniesieniami do określonych napędów. Mogło to przyjmować np. następującą strukturę:

- Dysk z systemem operacyjnym Z:
- Dysk z programem 1 Y:

.....

- | | |
|----------------------|----|
| • Dysk z programem 8 | R: |
| • CD-ROM 1 | Q: |
| • CD-ROM 2 | P: |
| ... | |
| • Dysk logujący | F: |

Zdecydowaną wadą takiego systemu, dostrzeganą już na pierwszy rzut oka, jest znacznie ograniczona ilość potencjalnych czytników - a co za tym idzie - również fizycznych dysków kompaktowych, które możemy udostępnić czytelnikom.

Generacja druga

Kolejnym etapem rozwoju technologicznego stało się integrowanie dużych zestawów czytników z systemami serwerowymi opartymi o produkty firm Novell i Microsoft, uzupełnionymi o wieże z czytnikami CD-ROM. Etap ten uzupełniony został specjalnie oprogramowaniem do ich zarządzania. Zasada działania oprogramowania opierała się na dynamicznym przydzielaniu liter, adresujących poszczególne czytniki CD-ROM w zależności od uruchomionej bazy danych. Przebieg tych procesów mógł być następujący:

BAZA PIERWSZA

- | | | | |
|----------|----|----------|----|
| - dysk 1 | K: | - dysk 2 | L: |
| - dysk 3 | M: | | |

BAZA DRUGA

- | | | | |
|----------|----|----------|----|
| - dysk 1 | K: | - dysk 2 | L: |
| - dysk 3 | M: | - dysk 4 | N: |

BAZA TRZECIA

- | | | | |
|----------|----|----------|----|
| - dysk 1 | K: | - dysk 2 | L: |
| - dysk 3 | M: | - dysk 4 | N: |
| - dysk 5 | O: | - dysk 6 | P: |
| - dysk 7 | Q: | - dysk 8 | R: |

Generacja trzecia

Główną wadą omówionych powyżej systemów było ograniczenie pracy tylko i wyłącznie do sieci lokalnych. Drugą ogromną wadą - utrudnienie równoczesnego dostępu kilku użytkowników do tego samego krążka CD. Jak się okazało, ciężkie głowice czytników CD-ROM, odczytując dane z płyt kompaktowych nie były zdolne do szybkiego reagowania na zapotrzebowanie użytkowników. W związku z powyższym koniecznością stało się m.in. przeładowywanie płyt CD na dyski twarde oraz stworzenie możliwości otwarcia dostępu do danych na świat. Zrealizowane zostało to w systemach InfoWare CD/HD instalowanych w całej Europie. W latach ubiegłych w Polsce użytkownikami ich stały się:

1. Uniwersytet Mikołaja Kopernika w Toruniu
2. Konsorcjum Górnośląskich Bibliotek Naukowych
3. Uniwersytet Warszawski w Warszawie
4. Akademia Górniczo-Hutnicza w Krakowie
5. Centrum Superkomputerowe w Poznaniu
6. Politechnika Gdańska w Gdańsku
7. Urząd Patentowy w Warszawie (system wykorzystujący jukebox'y)

Charakterystyczne cechy takich systemów to:

1. Oprogramowanie pracujące w oparciu o technologię klient-serwer z wykorzystaniem systemów wieloprotokołowych, głównie w opartych o TCP/IP i Netbios.
2. Bazy są przeładowywane na dyski twarde, co zapewnia odpowiednio szybki dostęp do danych dla wielu równoczesnych użytkowników tej samej bazy.
3. Użytkownik po zalogowaniu się do systemu otrzymuje menu, z którego wybiera żadaną pozycję - bazę danych lub jej element.
4. Bazy są przeglądane przy pomocy firmowego oprogramowania, dostarczanego wraz z nią przez jej producenta.
5. Podczas pierwszego uruchomienia bazy z serwera na dysk użytkownika jest w pobierane oprogramowanie przeszukujące, wykorzystywane później podczas następnych sesji wyszukiwania danych.
6. Logowanie i obsługa systemu wykonywana jest przy pomocy specjalnego klienta dostarczanego wraz z oprogramowaniem serwera.

Przykładem rozwiązania dla sieci rozległej jest Górnośląskie Konsorcjum Bibliotek Naukowych, gdzie użytkownikami systemu są:

- Jednostki organizacyjne Uniwersytetu Śląskiego i Akademii Ekonomicznej zlokalizowane w Katowicach
- Jednostki organizacyjne Uniwersytetu Śląskiego zlokalizowane w Sosnowcu
- Jednostki organizacyjne Uniwersytetu Śląskiego zlokalizowane w Cieszynie
- Jednostki organizacyjne Uniwersytetu Opolskiego zlokalizowane w Opolu

Jak się okazało, w toku użytkowania tychże systemów największe kłopoty wynikały z cech wymienionych w pkt. 5 i 6. Instalacja "klientów" i zaśmiecanie dysków użytkowników stały się kłopotliwe. Wady te zostały wyeliminowane w toku dalszego rozwoju technologicznego oprogramowania systemowego.

Generacja czwarta

Zmiany wymienione powyżej zostały przezwyciężone w systemach INFOWARE-IRIS. Systemy te przygotowano w dwóch mutacjach:

- CD WARE - system rozpowszechniania baz danych w sieciach lokalnych i szybkich sieciach metropolitalnych - wykorzystuje właściwości sieci Microsoft
- CD WEB WARE - system rozpowszechniania baz danych w sieciach lokalnych i szybkich sieciach metropolitalnych oraz wolnych np. modemowych - wykorzystuje serwery firmy CITRIX (Winframe i Metaframe)

Do cech charakterystycznych InfoWare Remote Information Server należą:

- możliwość modułowego rozwoju systemu poprzez współpracujące ze sobą serwery baz danych i serwery aplikacyjne
- możliwość wykorzystania praktycznie wszystkich platform sprzętowych i systemowych klienta (systemy z wykorzystaniem serwera aplikacyjnego), mimo iż oprogramowanie przeglądające bazę jest oferowane tylko w wersji dla Windows. Wdrażane rozwiązanie wykorzystuje DOS, Windows 95/98/NT, Unix we wszystkich mutacjach oraz produkty firmy Apple pracując poprzez protokół ICA.
- praca na wolnych łączach modemowych
- logowanie się do systemu poprzez aktualną wersję przeglądarki internetowej: Internet Eksplorator, Netscape-Sylaba lub specjalnie przygotowany IRIS Browser
- pobieranie i uruchamianie oprogramowania przeszukującego bazy danych bezpośrednio z serwera (bez "zaśmiecania" dysków użytkowników kolejnymi jego wersjami)
- możliwość integracji w systemie praktycznie wszystkich baz danych dostępnych na naszym rynku, dystrybuowanych na nośnikach optycznych
- zabezpieczenie licencyjnego użytkownika systemu poprzez kontrolę nr IP użytkownika

- udostępnianie katalogów i pozycji menu na podstawie systemu nazw użytkowników

Najważniejsza różnica pomiędzy wymienionymi powyżej dwoma wersjami oprogramowania IRIS polega na tym, że gdy pierwsze całość operacji wykonuje na komputerze czytelnika, to wersja druga realizuje przeszukiwania przy pomocy serwera aplikacyjnego, przekazując do czytelnika tylko wygląd ekranu poprzez kompresujący protokół ICA (dla przekazania całości obrazu o rozdzielczości 800*600 przy 256 kolorach potrzeba tylko około 5 kB danych)

Systemy te pracują w Polsce obsługując:

- Konsorcjum Bibliotek CDROMWAN
- Górnośląskie Konsorcjum Bibliotek Naukowych
- Szkołę Główną Gospodarstwa Wiejskiego
- Politechnikę Wrocławską na potrzeby uczelni Wrocławia

Ogromnie ważną cechą tego oprogramowania jest rozbudowany moduł rejestracji logowań do systemu. Analizy można prowadzić pod różnymi aspektami (np. czasu pracy poszczególnych użytkowników, wykorzystania poszczególnych baz danych, ilości pobieranych danych, ilości błędnych uruchomień itd.), jak również dane te mogą zostać poddane dalszej obróbce, po wyprowadzeniu ich do plików w kilku popularnych formatach (np. *.xls, *.rtf, *.txt). Biblioteka Uniwersytetu Śląskiego w oparciu o ten system przygotowała serwer dostępowy dla wybranych jednostek naukowych do baz cytowań Science Citation Index, Social Sciences Citation Index oraz Art And Humanities Citation Index, opracowanych przez Institute for Scientific Information

System ERL firmy Silver Platter

System ten składa się z kilku modułów, które łącząc swoje możliwości tworzą całokształt cech użytkowych systemu. Należą do nich:

- Serwer ERL z modułem administracyjnym
- Oprogramowanie przeszukujące bazy danych WebSpirs
- Oprogramowanie przeszukujące bazy danych Winspirs, PCSpirs, MacSpirs

Podstawowym elementem sieciowego rozpowszechniania baz danych firmy Silver Platter jest serwer ERL. W zależności od potrzeb - ilości równoczesnych potencjalnych lub rzeczywistych użytkowników oraz ilości zainstalowanych baz danych -- system ten może zostać uruchomiony na komputerach o różnych mocach obliczeniowych, z zainstalowanymi różnymi systemami operacyjnymi. Firma Silver Platter produkuje oprogramowanie serwera w następujących wersjach:

- ERL dbserver 4.05 for AIX
- ERL dbserver 4.05 for Red Hat Linux v6
- ERL dbserver 4.05 for SCO UNIX
- ERL dbserver 4.05 for SUN Solaris
- ERL dbserver 4.05 for Windows NT

System ten po uruchomieniu na dowolnej platformie systemowej umożliwia przeszukiwanie zainstalowanych tam baz danych, przygotowanych wcześniej w strukturze firmy Silver Platter. Do uruchomienia takiej bazy danych niezbędne jest wcześniejsze wprowadzenie przekazywanych przez firmę Silver Platter kodów, w których zawarte są m.in.:

- Numer licencji
- Nazwa bazy danych
- Nazwa użytkownika wraz z jego siedzibą

- Identyfikator serwera
- Identyfikator bazy danych
- Maksymalna ilość użytkowników
- Parametry instalacji bazy (m.in. możliwość przeładowywania na dysk twardy)
- Kod kontrolny

Definicje wymienionych parametrów umożliwiają pełną kontrolę dotyczącą sieciowego wykorzystania baz danych.

Oprogramowanie ERL pracuje jako system w wersji klient-serwer. Klientem do serwera ERL są wszystkie mutacje oprogramowania do przeszukiwania baz danych. System może pracować w następujących wersjach:

- Klient - oprogramowanie WinSpirs
Po zadeklarowaniu parametrów serwera ERL (numeru IP) i uruchomieniu oprogramowania następuje nawiązanie łączności z serwerem, zalogowanie do systemu (może odbywać się automatycznie na podstawie wcześniej zapisanych na dysku twardym parametrów) oraz otwarcie tablicy dostępnych baz danych. W przypadku zadeklarowania większej ilości serwerów ERL, logowanie do nich wszystkich następuje równocześnie. Podczas logowania do systemu następuje autoryzacja wykonywana na podstawie nr IP i nazwy użytkownika. W związku z możliwością wybrania różnych baz danych (nawet z niepokrewnych dziedzin), przeszukiwanie baz danych można znacznie uprościć. Identyczne działanie mają również oprogramowanie MacSpirs.
- Tryb poprzez przeglądarkę internetową
Podczas konfiguracji systemu na serwerze WWW instalujemy oprogramowanie WebSpirs. Oprogramowanie to jest przystosowane do współpracy z dowolną przeglądarką internetową. Poprzez nadany jej systemowy odnośnik - z jednej strony przyjmuje zapytania od klienta wysyłającego je z formatu HTML, przekazując do serwera ERL - a z drugiej, po otrzymaniu odpowiedzi od niego, wysyła je w formacie HTML do pytającego wcześniej klienta. Również i w tym wypadku prowadzona jest kontrola uprawnień klienta poprzez jego numer IP.
- Tryb mieszany
Jak na specjalistyczny system przystało możliwa jest również praca mieszana. I tak: wszyscy użytkownicy podłączeni w ramach sieci lokalnych, a nawet szybkich światłowodowych sieci miejskich, mogą korzystać z posadowionych baz danych na serwerze ERL pracując w WinSpirs'ie, natomiast wszyscy ci, którzy nie mają szybkich łącz, korzystają z nich poprzez przeglądarkę internetową.

Ważne ze strony użytkownika są również wszelkie funkcje dodatkowe oferowane przez systemy Silver Platter. Należą do nich m.in.:

1. Możliwość przeszukiwania zapomocą hierarchicznego systemu haseł tematycznych
2. Możliwość "mapowania", tj. przydzielania dowolnego słowa do terminu indeksowego zaczerpniętego z tezausa
3. Możliwości linkowania:
 - o system SilverLinker, który już umożliwia linkowanie rekordów bibliograficznych do pełnych tekstów (800 000 dokumentów, 450 000 odrębnych artykułów w 1902 tytułach czasopism); współpraca z 12 dostawcami czasopism w wersji elektronicznej, w tym do EBSCO za darmo (dzięki dofinansowaniu SOROS-a)
 - o możliwość linkowania rekordów bibliograficznych do odpowiadających im tekstów w pełnoobrazowych bazach wydawców np. ELSEVIER
4. Współpraca z bazami bibliograficznymi bibliotek:
 - o system SilverLinker, który umożliwia natychmiastowe linkowanie rekordów bibliograficznych do odpowiadających im rekordów w dowolnej liczbie systemów OPAC wykorzystując tzw. "żywe" bazy danych
 - o możliwość linkowania rekordów bibliograficznych do zamkniętych baz katalogowych rozprowadzanych na CD-ROM'ach
5. Pełna możliwość wykorzystania:
 - o SDI (użytkowników) - ustalania własnego profilu użytkownika do

- wielokrotnych następujących w odstępach czasowych przeszukiwań
 - alertów (tj. SDI zarządzanych przez bibliotekarzy)
 - zarządzania obsługą wypożyczeń międzybibliotecznych (ILL)
- 6. Stosunkowo proste generowanie wieloaspektowych statystyk użytkowania, z natychmiastowym podglądem lub / i przeładowaniem do formatu np. Excel'a (*.xls).
- 7. Możliwość lokalnego zaprojektowania systemu menu - dowolny wygląd i język.
- 8. Możliwość wyszukiwania wielobazowego (narusz) zgodnie z deklaracją na początku sesji.
- 9. Możliwość zawężania wyszukiwania do wskazanych przedziałów chronologicznych (np. tylko okres od 1995-1997)
- 10. Obszerne materiały szkoleniowe dostępne on-line
 - Manual online
 - Tutorial online - www.silverplatter.com/tutor
 - Tutorial for INSPEC on SilverPlatter (powstał ze współpracą z INTELEM)

Omówienie pozostałych systemów, nie mających szerszej reprezentacji w Polsce, w następnym odcinku.
